JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 44, No. 4 AMERICAN WATER RESOURCES ASSOCIATION August 2008

SPATIAL CALIBRATION AND TEMPORAL VALIDATION OF FLOW FOR REGIONAL SCALE HYDROLOGIC MODELING[1]

C. Santhi, N. Kannan, J. G. Arnold, and M. Di Luzio[2]

ABSTRACT: Physically based regional scale hydrologic modeling is gaining importance for planning and management of water resources. Calibration and validation of such regional scale model is necessary before applying it for scenario assessment. However, in most regional scale hydrologic modeling, flow validation is performed at the river basin outlet without accounting for spatial variations in hydrological parameters within the subunits. In this study, we calibrated the model to capture the spatial variations in runoff at subwatershed level to assure local water balance, and validated the streamflow at key gaging stations along the river to assure temporal variability. Ohio and Arkansas-White-Red River Basins of the United States were modeled using Soil and Water Assessment Tool (SWAT) for the period from 1961 to 1990. R^2 values of average annual runoff at subwatersheds were 0.78 and 0.99 for the Ohio and Arkansas Basins. Observed and simulated annual and monthly streamflow from 1961 to 1990 is used for temporal validation at the gages. R^2 values estimated were greater than 0.6. In summary, spatially distributed calibration at subwatersheds and temporal validation at the stream gages accounted for the spatial and temporal hydrological patterns reasonably well in the two river basins. This study highlights the importance of spatially distributed calibration and validation in large river basins.

(KEY TERMS: spatially distributed calibration; validation; hydrologic modeling; regional scale; HUMUS; SWAT; CEAP.)

Santhi, C., N. Kannan, J.G. Arnold, and M. Di Luzio, 2008. Spatial Calibration and Temporal Validation of Flow for Regional Scale Hydrologic Modeling. *Journal of the American Water Resources Association* (JAWRA) 44(4):829-846. DOI: 10.1111/j.1752-1688.2008.00207.x

INTRODUCTION

There are serious concerns about managing the water quantity and quality throughout the United States (U.S.) (USEPA, 1998). As water resource systems often cross local and state boundaries, the planning and management processes often require a basin-wide or regional perspective. Compared to the traditional approach of looking at a specific watershed, a regional planning approach can help to develop a comprehensive vision for future growth, and develop plans to use and manage the water resources efficiently. However, management and utilization of water resources in a region depends upon the spatial and temporal distribution of rainfall,

[1]Paper No. J06179 of the *Journal of the American Water Resources Association* (JAWRA). Received December 16, 2006; accepted February 20, 2008. **Discussions are open until February 1, 2009**.

[2]Respectively (Santhi, Kannan, Di Luzio) (Associate Research Scientist, Assistant Research Scientist, Research Scientist), Blackland Research and Extension Center, Texas A&M University System, 720 East Blackland Road, Temple, Texas 76502; and Supervisory Agricultural Engineer and Research Leader, Grassland Soil and Water Research Laboratory, Agricultural Research Service-U.S. Department of Agriculture, 808 East Blackland Road, Temple, Texas 76502 (E-Mail/Santhi: csanthi@brc.tamus.edu).

runoff, ground-water storage, evapotranspiration (ET), soil types and crops grown. These factors vary from basin to basin or region to region. Therefore, understanding and capturing the spatial and temporal variability of these factors on hydrological pattern both at subwatershed and watershed levels is necessary.

Physically based regional scale hydrologic modeling (with geographic information system [GIS] capability) can simulate the spatial and temporal variability of hydrological processes in different subunits of the region. It can be used for investigating the impacts of different water quality management alternatives in different subunits and develop management plans. However, development of regional scale models is a difficult task, because of the spatial and temporal scales that must be considered and the large amount of information that must be integrated. The other difficult task is the calibration of the model at regional scale. Only limited attempts had been made to develop, apply, and validate physically based hydrological models for regional scale studies.

Jha *et al.* (2006) have used physically based model, Soil and Water Assessment Tool (SWAT) in combination with General Circulation Model for regional scale climate studies in the U.S. Hao *et al.* (2004) modeled the Yellow River Basin in China and calibrated and validated the flow. In most regional or large-scale modeling studies including the above, simulated flow is calibrated and validated against measured streamflow at one or two gaging stations on the river mostly at the watershed outlet. This is accomplished by adjusting the model inputs for the entire watershed to match the flows at the selected gages without adequate validation in various subunits or subwatersheds of the region. One of the major limiting factors for this is the availability of observed flow data for calibration and validation. However, it is important to note that there are wide variations in runoff produced in different subunits of the large river basin due to variations in rainfall, soils, land use and vegetation and the associated hydrological processes. It is necessary to capture the spatial and temporal variability of hydrologic pattern across the region with adequate calibration of runoff in different subunits (Arnold *et al.*, 2000) and also the temporal variations of flow patterns. Runoff being an important component of the water balance, capturing the variation in runoff will represent the hydrologic pattern in the watershed. Qi and Grunwald (2005) have calibrated and validated the simulated flow against measured stream flow in the Sandusky watershed at the watershed outlet and four other subwatershed locations using SWAT. Their approach captured the spatial and temporal variations in flow in the four subwatersheds and the watershed. Their study was conducted on a relatively small watershed (drainage area 3,240 km^2) and calibration and validation was conducted for two years each.

The calibration and validation approach used in this study is different from the above studies. The specific objectives of this study are to conduct:

(1) a spatially distributed calibration of long-term average annual runoff at subwatershed level in a regional scale river basin to capture the spatial variation in runoff in different parts of the river basin, and
(2) a temporal validation of streamflow at key locations (gage) along the river.

The spatial calibration helps in assuring local water balance at subwatershed level. The temporal validation is performed to assure annual and seasonal variability. It is expected that the calibration and validation approach used for large river basins in this study would improve the reliability of hydrologic model predictions in regional scale river basins and also would improve our knowledge of local hydrological patterns nested within a large basin. This approach would also be useful for modelers, researchers and planners involved in regional scale studies.

This study was conducted as part of an on-going national scale assessment study, Conservation Effects Assessment Project (CEAP). CEAP follows the HUMUS/SWAT (Hydrologic Unit Modeling for the U.S.) watershed modeling framework (Srinivasan *et al.*, 1998). Within the HUMUS framework each water resource region (major river basin) is treated as a watershed and each U.S. Geological Survey (USGS) delineated eight-digit watershed as a sub-watershed for use in SWAT modeling. The HUMUS system is updated with recently available databases and SWAT model for the CEAP—National Assessment (Santhi *et al.*, 2005; Di Luzio *et al.*, 2008). The main objective of the CEAP study is to quantify the environmental and economic benefits obtained from the conservation practices and programs implemented in the U.S. The benefits are reported at the eight-digit watershed and river basin scales. Therefore, the hydrologic model used for such assessment is expected to simulate the flow and pollutant transfer reasonably well in all the eight-digit watersheds and time series of streamflow at key locations along the main river system. The regional scale calibration and validation approach described in this study is used for CEAP. This paper describes the regional scale hydrologic modeling and spatial and temporal calibration and validation of flow in two river basins with different hydrologic conditions (a high flow and a low flow region based on annual average rainfall and runoff).

METHODOLOGY

The CEAP/HUMUS system used in this study consists of a hydrologic/watershed scale model, SWAT (Arnold *et al.*, 1998; Neitsch *et al.*, 2002; http://www.brc.tamus.edu/swat) and revised databases for preparing the model inputs (Santhi *et al.*, 2005; Di Luzio *et al.*, 2008). SWAT was selected because of its ability to simulate land management processes in large watersheds. SWAT has been widely used in the U.S. and other countries (Arnold *et al.*, 1999; Borah and Bera, 2004; Gassman *et al.*, 2007). Arnold *et al.* (1999) and Gassman *et al.* (2007) have reported previous model validation studies of several locations throughout the U.S. Borah and Bera (2004) have extensively reviewed various models and indicated that SWAT is a suitable model for long-term continuous simulations of large watersheds.

SWAT Model Description

SWAT is a physically based, semi-distributed model developed to simulate continuous-time landscape processes and streamflow with a high level of spatial detail by allowing the river/watershed to be divided into a large number of subbasins or subwatersheds. Each subbasin is further divided into several unique land use and soil combinations called Hydrologic Response Units (HRUs) based on threshold percentages used to classify the land use and soil (Arnold *et al.*, 1998; Neitsch *et al.*, 2002) and they are homogeneous. SWAT operates on a daily time step and is designed to simulate water, sediment and agricultural chemical transport in a large ungaged basin and evaluate the effects of different management scenarios on watershed hydrology and point and nonpoint source pollution. Key components of the model include hydrology, weather, erosion, soil temperature, crop growth, nutrients, pesticides, and agricultural management. A complete description of all components can be found in Arnold *et al.* (1998) and Neitsch *et al.* (2002). A brief description on flow is provided here.

Upland Processes/Hydrology. The local water balance in the Hydrologic Response Unit is provided by four storage volumes: Snow (stored volume until it melts), soil profile (typically 0-2 m), shallow aquifer (typically 2-20 m), and deep aquifer (>20 m). The soil profile can be subdivided into multiple layers. Soil water processes include infiltration, runoff, evaporation, plant uptake, lateral flow, and percolation to lower layers. Percolation from the bottom of the soil profile recharges the shallow aquifer (ground-water recharge). Surface runoff from daily rainfall is estimated with a modification of U.S. Department of Agriculture-Soil Conservation Service (SCS) curve number method (USDA-SCS, 1972). Green & Ampt infiltration method is also available within SWAT to simulate surface runoff and infiltration. SWAT has options to estimate the potential evapotranspiration (PET) by different methods such as Modified Penman Montieth, Hargreaves, and Priestley-Taylor. The Hargreaves method is used in this study (Hargreaves *et al.*, 1985). Flow generation, sediment yield and nonpoint source loadings from each HRU in a subwatershed are summed and the resulting flow and pollutant loads are routed through channels, reservoirs and/or ponds to the watershed outlet.

Channel Processes. Channel processes simulated within SWAT include flood routing, sediment routing and nutrient and pesticide routing. Ponds/Reservoirs components including water balance, routing, sediment settling and simplified nutrient and pesticide transformation are used in SWAT (Neitsch *et al.*, 2002).

Study Area Description

Two river basins or water resources regions with different climatic conditions, runoff, land use distribution, vegetation, soils, and topography have been modeled to capture the spatial and temporal variations involved in the hydrologic processes and demonstrate the validity of the regional/basin scale modeling effort. The two regions studied are as follows: (1) The Ohio River Basin located in the eastern U.S., and (2) The Arkansas-White-Red River Basin located in the south central U.S. (Figure 1). These two regions are also referred to as Region 05 and Region 11 by the USGS at a 2-digit watershed scale or hydrologic accounting unit.

Ohio River Basin. The Ohio River starts at the confluence of the Allegheny and the Monongahela in Pittsburgh, Pennsylvania, and ends in Cairo, Illinois, where it flows into the Mississippi River. It flows through six states: Pennsylvania, West Virginia, Ohio, Illinois, Indiana and Kentucky (Figure 1). There are several dams across the Ohio River including many lock and dams built to facilitate navigation. The region is comprised of 120 USGS delineated eight-digit watersheds. The Ohio River Basin receives a high amount of rainfall. Agriculture is the predominant land use in this region and about 21% of the land is used as cropland (Table 1). The predominant soils in the region are gilpin, hazleto, zanesview,

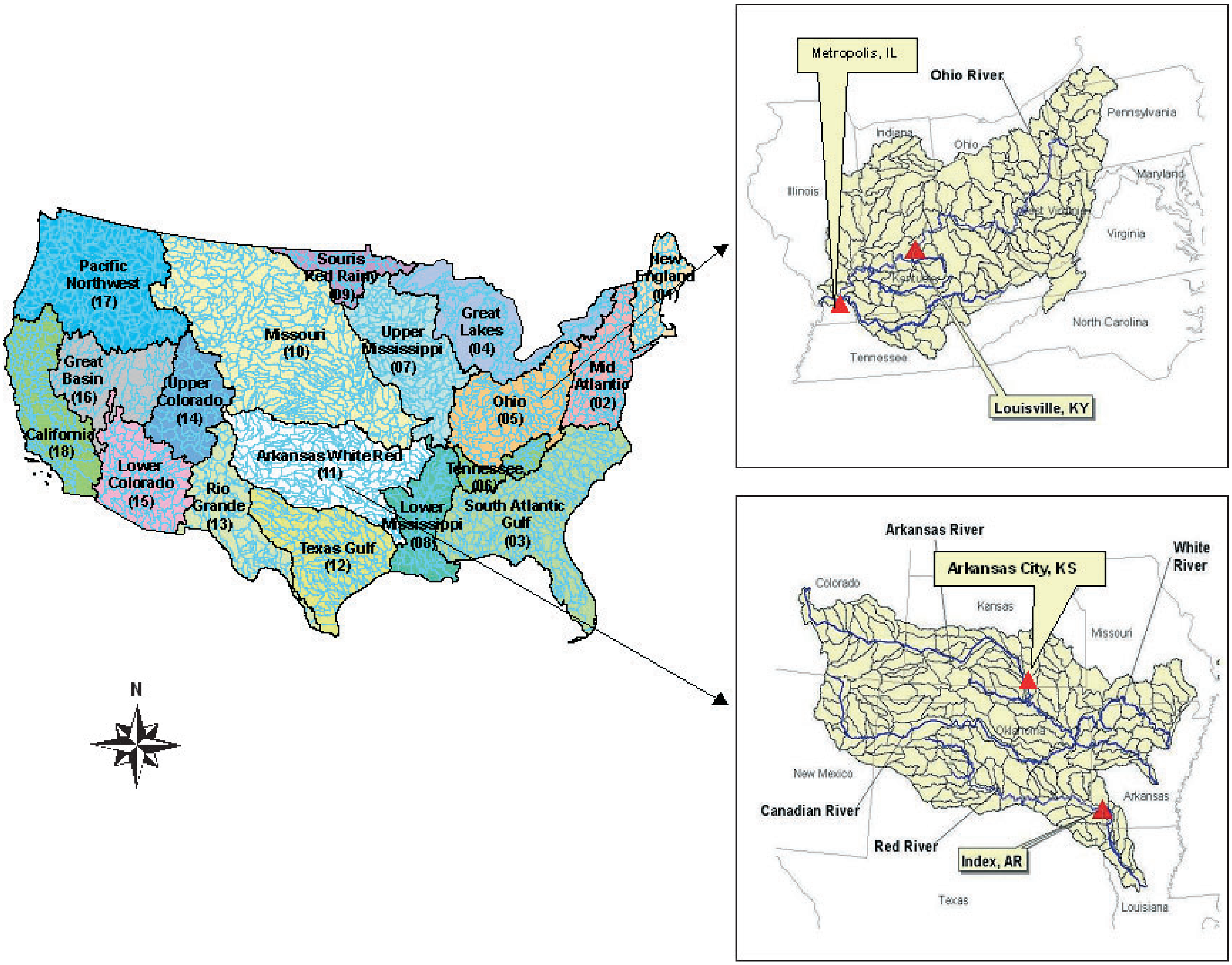

FIGURE 1. The Ohio and Arkansas-White-Red River Basins in the Conterminous United States With Flow Validation Gaging Stations.

TABLE 1. Watershed Characteristics of Ohio and Arkansas-White-Red River Basins.

River Basins/Regions	Ohio River Basin (Region 05)	Arkansas-White-Red River Basin (Region 11)
Number of eight-digit watersheds	120	173
Average annual rainfall at eight-digit watersheds* (mm)	1,140	800
Estimated average annual PET and ET at eight-digit watersheds* (mm)	1,100 and 700	1,500 and 600
Estimated average annual runoff at eight-digit watersheds* (mm)	440	150
Predominant land uses (%)	Forest (51%), Pasture/Hay (22%), Cropland (21%), Urban (3%) and others (3%)	Range (41%), Forest (22%), Pasture/Hay (19%), Cropland (14%) and others (4%)

*Average values of 30 years, approximated or rounded off.

faywood, and bodine (USDA-NRCS, 1994). This region has also witnessed increased population growth, urbanization and industrial development. Nonpoint source pollution from agricultural activities and urban runoff are major sources of pollution in this river basin.

Arkansas-White-Red River Basin. The drainage area of this region includes (1) the Arkansas River, within and between the States of Colorado, Kansas, Oklahoma, and Arkansas; (2) the White River, within and between the States of Missouri and Arkansas; and (3) the Red River, within and between the States of New Mexico, Texas, and Louisiana (Figure 1). These major rivers flow generally from west to east. The region is comprised of 173, eight-digit watersheds. Although most of the regions are dominated by range and forest, about 14% of the land is used for growing crops (Table 1). Dominant soils in the region are carnasa, stephen, berda, clarksv, dunipha, enders, pullman, and manvel (USDA-NRCS, 1994). The prevalent water quality problem in this region is from nonpoint source pollution. Sediment and nutrients are the major causes of nonpoint source pollution, especially in the State of Oklahoma. Some of the largest poultry and swine operations in the U.S. are located in this region. The impacts have become more prevalent in the streams, rivers, and lakes. Many of these lakes are the major drinking water source for large cities in this region.

Databases and Model Inputs

The HUMUS/SWAT system requires several data such as land use, soils, management practices, weather, point source data, and reservoirs. Considerable effort has been made to process and update the HUMUS/SWAT databases for CEAP and prepare SWAT input files for the river basins (Santhi *et al.*, 2005; Di Luzio *et al.*, 2008). The various databases used are described here.

Land Use. The 1992 USGS—National Land Cover Dataset at 30 m resolution was used in this study and it included land use classes such as cropland (row/small grains), urban, pasture, range, forest, wetland, barren, and water. Land use-related information is input to the model at HRU level.

Soils. Each land use within a subbasin is associated with soil data. Soil data required for SWAT were processed from the State Soil Geographic (STATSGO) database (USDA-NRCS, 1994). Each STATSGO polygon contains multiple soil series and the aerial percentage of each soil series. The soil series with the largest area was extracted and the associated physical properties of the soil series were used. Soil properties used in modeling include texture, bulk density, saturated hydraulic conductivity, available water holding capacity (AWC), total depth of soil, and organic carbon. Soil information is input for each HRU.

Management Data. Management operations, such as planting, harvesting, applications of fertilizers, manure, and pesticides and irrigation water and tillage operations are used for various land uses in the management files. A crop parameter database available within SWAT (Neitsch *et al.*, 2002) is used to characterize and simulate the crop growth defined in the management file. Management information is input at HRU level.

Topography. Elevation information is used to delineate the watersheds into different subwatersheds. Accumulated drainage area, overland field slope, overland field length, channel dimensions, channel slope, and channel length are derived for each subwatershed using the 3-arc second digital elevation model (DEM) data (Srinivasan *et al.*, 1998). Information extracted from DEM is used at subwatershed and HRU levels.

Weather. Measured daily precipitation and maximum and minimum temperature data from 1960 to 2001 are used in this study. The precipitation and temperature datasets are newly created (Di Luzio *et al.*, 2008) from a combination of point measurements of daily precipitation and temperature (maximum and minimum) (Eischeid *et al.*, 2000) and PRISM (Parameter-elevation Regressions on Independent Slopes Model) (Daly *et al.*, 2002). The point measurements compose serially complete (without missing values) dataset processed from the station records of the National Climatic Data Center. PRISM is an analytical model that uses point data and a digital elevation model to generate gridded estimates of monthly climatic parameters and distributed at 4 km^2. Di Luzio *et al.* (2008) have developed a novel approach to combine the point measurements and the monthly PRSIM grids to develop the distribution of the daily records with orographic adjustments over each of the USGS eight-digit watersheds. Other data such as solar radiation, wind speed and relative humidity are simulated using the monthly weather generator parameters from weather stations (Nicks, 1974; Sharpley and Williams, 1990) available within SWAT database for these regions. Weather data are input for each subwatershed.

Point Source Data. Effluents discharged from the municipal treatment plants are major point sources of pollution. The USGS has developed a point source database for use in the SPARROW (SPAtially Referenced Regressions On Watershed attributes) model simulations (Smith *et al.*, 1997) and it is used in this study. Point source data used include effluent discharge/flow and sediment and nutrient loadings. Point source data are input at subwatershed level.

The point source data was updated for 2000 population.

Reservoirs. Basic reservoir data such as storage capacity and surface area were obtained from the dams database (U.S. Army Corps of Engineers, 1982; Hitt, 1985). Because of the lack of adequate reservoir release data and complexity involved in simulating each reservoir operation in large-scale modeling effort, a simple reservoir simulation approach available within the SWAT model is used with a monthly target release-storage approach based on the storage capacity and flood and nonflood seasons (Neitsch *et al.*, 2002). Reservoir related data are input at the subwatershed level.

Calibration and Validation Approach Used for Regional Scale Hydrologic Modeling

For this regional scale modeling, a calibration procedure involving (1) calibration of spatial variations in runoff at subwatershed level, and (2) temporal validation of streamflow at multiple gaging stations along the major river, was used. The model was run using weather data from 1960 through 1990 for the two river basins studied. Data from 1960 is used for the model to assume realistic initial conditions and it was not included in the calibration. Data from 1961-1990 is used for calibration. Subwatersheds and eight-digit watersheds are used interchangeably in this paper.

Spatially Distributed Calibration. In general, spatial calibration refers to the calibration of a watershed model with known "spatially distributed" input and output information. SWAT is a physically based, semi-distributed watershed model and the watershed is disaggregated in geographical space and their processes in time. Thus, analysis can be described in a spatial and temporal context. Subwatershed is considered as the "spatial unit of variation" for this regional scale calibration study because this is a relatively large area study to meet the needs of the CEAP national assessment. From a watershed modeling perspective, SWAT is capable of simulating a high level of spatial detail by allowing a watershed to be divided into multiple subwatersheds. Heterogeneity in inputs within a subwatershed is captured by dividing the subwatershed into several HRUs, which are unique land use soil combinations. More land use and soil combinations (more HRUs for increased spatial detail) within a subwatershed can be obtained by using the lowest threshold level for selecting land use and soil combinations.

Observed data used for spatially distributed calibration: Gebert *et al.* (1987) has prepared the average annual runoff contour map for the conterminous United States using measured streamflow data from 5,951 USGS gaging stations for the period 1951-1980 and the stream flows at these gauging stations were considered to be natural (i.e., unaffected by upstream reservoirs or diversions) and representative of local conditions. For this study, the contours of average annual runoff produced by Gebert *et al.* (1987) were interpolated to produce a smooth grid using GIS interpolation technique called inverse distance weighting and the average annual runoff for eight-digit watersheds were estimated. Although the runoff estimated is long-term annual average, it was still used for calibration to capture the spatial variations in runoff at subwatershed level because of a large (regional) study area, adequacy of the project needs, and limitation in availability of time series observed data. The average annual runoff is a good indicator of annual water balance in a subwatershed or watershed although it may not readily convey the temporal variation effects and seasonality effects. Several studies have used the average runoff contours for regional scale studies and showed the spatial variation in runoff across the region (Wolock and Mc Cabe, 1999). Hence, it is important to capture the spatial variation in runoff during calibration.

Calibration Parameters: The model is calibrated to capture the spatial variation in long-term average annual observed runoff by adjusting several model input parameters (Table 2), keeping them within realistic uncertainty ranges. Calibration is performed for each subwatershed (eight-digit) by adjusting the model parameters that capture the spatial and temporal variations in model inputs such as soils, land use, topography, and weather and interactions among them influencing various hydrologic processes, such as runoff, ET, and ground-water flow. The input parameters used for calibration (Table 2) include

(1) HARG_PETCO is a coefficient used to adjust potential evapotranspiration (PET) estimated by Hargreaves method (Hargreaves and Samani, 1985; Hargreaves and Allen, 2003) and calibrate the runoff in each subwatershed. In Hargreaves method, PET is related to temperature and terrestrial radiation. This coefficient is related to radiation and can be varied to match the PET in different parts of the region depending on the weather conditions (Hargreaves and Allen, 2003).
(2) Soil water depletion coefficient (CN_COEF) is a coefficient used in the curve number method to adjust the antecedent moisture conditions on surface runoff. This parameter is related to

TABLE 2. Input Parameters Used in the Calibration Procedure, Their Range and Their Effects on Different Components of Runoff.

Parameter	Description	Spatial Level of Parameterization	Changes: Surface Runoff	Changes: Ground Water	Changes: Water Yield	Range Used: Min	Range Used: Max
HARG_PETCO	Coefficient used to adjust potential evapotranspiration estimated by Hargreaves method and runoff	Subwatershed (HRU)	X	X	X	0.0019	0.0027
Soil Water Depletion Coefficient (CN_COEF)	Coefficient used in the new curve number method (Neitsch *et al.*, 2002; Kannan *et al.*, 2008) and is used to adjust surface runoff and groundwater in accordance with soil water depletion.	Subwatershed (each soil in HRU)	X	X	X	0.5	1.50
CN	Curve number—adjust surface runoff	HRU (for each landuse and soil)	X		X	−5	+5
GWQMN	Minimum threshold depth of water required in shallow aquifer for ground-water flow to occur	Subwatershed		X	X	−3	+3
GWREVAP	Groundwater re-evaporation coefficient that controls the upward movement of water from shallow aquifer to root zone in proportion to evaporative demand	Subwatershed		X	X	0.02	0.20
AWC	Soil available water holding capacity	HRU		X	X	−0.04	+0.04
ESCO	Soil evaporation compensation factor, that is used to modify the depth distribution of water in soil layers to meet the soil evaporative demand)	HRU		X	X	0.50	0.99
EPCO	Plant evaporation compensation factor, that allows water from lower soil layers to meet the potential water uptake in upper soil layers	HRU		X	X	0.01	0.99

PET, precipitation and runoff in the curve number method.

(3) Curve number (CN) is used to adjust surface runoff and relates to soil and land use and hydrologic condition at HRU level.

(4) Ground-water re-evaporation coefficient (GWREVAP) controls the upward movement of water from shallow aquifer to root zone, due to water deficiencies, in proportion to PET. This parameter can be varied depending on the land use/crop. The revap process is significant in areas where deep rooted plants are growing and affects the groundwater and the water balance.

(5) GWQMN-minimum threshold depth of water in the shallow aquifer to be maintained for ground-water flow to occur to the main channel.

(6) Soil AWC, which varies by soil at HRU level.

(7) Soil evaporation compensation factor (ESCO), which controls the depth distribution of water in soil layers to meet soil evaporative demand. This parameter varies by soil at HRU level.

(8) Plant evaporation compensation factor (EPCO), that allows water from lower soil layers to meet the potential water uptake in upper soil layers and varies by soil at HRU level.

The input parameters were adjusted within literature reported ranges (Santhi *et al.*, 2001; Neitsch *et al.*, 2002). Additional details of these parameters can be found in Neitsch *et al.* (2002). Effects of these input parameters on different components of runoff are shown in Table 2. It should be noted that an adjustment in runoff (due to changes in model parameters) results in changes in surface runoff and/or groundwater. Similarly, changes in surface runoff and groundwater result in changes in runoff.

The calibration process for each eight-digit watershed is carried out in three steps viz. (1) calibration of runoff (by adjusting HARG_PETCO), (2) surface runoff (by adjusting soil water depletion-coefficient and curve number), and (3) ground-water (all the other parameters mentioned in Table 2). An automated procedure is developed for conducting the spatially distributed calibration process at eight-digit watersheds in the river basin (Kannan *et al.*, 2008). Simulated runoff in each eight-digit watershed was calibrated by adjusting the model input parameters until average annual observed runoff and simulated runoff were within 20%. In this study, the simulated runoff (same as water yield) is defined as the sum of surface runoff, lateral flow from the soil profile and ground-water flow from the shallow aquifer simulated by the model. It is expected that this spatial calibration procedure (with minimal or no additional calibration) can provide good results in predictions of annual and monthly flows.

Temporal Validation of Flow at Multiple Gages Along the Main River. The temporal validation approach is useful in evaluating the model performance during high and low flow years, annual, seasonal, and monthly variations and in understanding the long-term temporal variations in hydrologic processes. Such a long-term study is necessary for planning and implementing conservation measures and programs and evaluating their performance. It should be noted that minimal attempt was made to adjust the model parameters or do additional calibration during temporal flow validation.

Observed data used for temporal validation: Annual and monthly streamflow data from USGS gaging stations at key locations along the main river representing different drainage area were selected and used to validate the simulated flow to assure proper annual and seasonal variability (Figure 1 and Table 3).

Statistical measures used for model evaluation: Several statistical measures including mean, standard deviation, coefficient of determination (R^2) and Nash-Suttcliffe efficiency (NSE) (Nash and Suttcliffe, 1970) were used to evaluate the annual and monthly simulated flows against the measured flows at the gages. If the R^2 and NSE values are less than or very close to 0.0, the model prediction is considered "unacceptable or poor." If the values are 1.0, then the model prediction is considered "perfect." A value greater than 0.6 for R^2 and a value greater than 0.5 for NSE, were considered acceptable (Santhi *et al.*, 2001; Moriasi *et al.*, 2007).

TABLE 3. Summary of Validation Results for Gaging Stations in the Ohio and Arkansas-White-Red River Basins.

	Name	Ohio	Ark-White-Red
Region	**Region**	**Region 05**	**Region 11**
Gage details	River	Ohio	Red
	Location	Louisville, KY	Index, AR
	Station ID	03294500	07337000
	Drain Area (km^2)	236,130	124,398
Annual	Mean (O) mm	451	87
	Mean (S) mm	433	75
	StdDev (O) mm	100	45
	StdDev (S) mm	82	42
	R^2	0.94	0.86
	NSE	0.86	0.79
Monthly	Mean (O) mm	38	7
	Mean (S) mm	36	6
	StdDev (O) mm	28	8
	StdDev (S) mm	20	7
	R^2	0.83	0.66
	NSE	0.72	0.64
Gage details	River	Ohio	Arkansas
	Location	Metropolis, IL	Arkansas City, KS
	Station ID	03611500	07146500
	Drain Area (km^2)	525,770	113,217
Annual	Mean (O) mm	491	14
	Mean (S) mm	467	10
	StdDev (O) mm	122	6
	StdDev (S) mm	100	7
	R^2	0.95	0.71
	NSE	0.89	0.13
Monthly	Mean (O) mm	41	1.3
	Mean (S) mm	39	1.0
	StdDev (O) mm	28	1.4
	StdDev (S) mm	22	2.0
	R^2	0.83	0.64
	NSE	0.81	0.23

Notes: O, observed; S, simulated; StdDev, standard deviation.

RESULTS AND DISCUSSION

Calibration of Spatial Variation in Runoff

Precipitation, simulated ET, simulated runoff, and observed runoff for the Ohio and Arkansas regions show the variations in hydrological patterns across eight-digit watersheds (Figures 2 and 3). It was noted that the observed runoff estimated at eight-digit watersheds varied widely ranging from <200 mm through 750 mm in the Ohio River Basin and it varied from <50 mm through 530 mm in the Arkansas River Basin. Hence, it is important to account for this spatial variation in runoff across the subwatersheds as opposed to traditionally calibrating the model inputs over the entire basin to match the flow at one stream gage at the watershed outlet. In order to illustrate how some of the model input parameters spatially affect the simulated outputs (either runoff or surface runoff or ground water), three such model input parameters are discussed here. Depending on the type of land use, crops grown, soil types, precipitation, and evaporation in each subwatershed, effects of these parameters on simulated hydrology (runoff, surface runoff, and ground water and ET) varied across subwatersheds.

1. HARG_PETCO, is used to calibrate the runoff by adjusting the PET within each subwatershed. PET is related to maximum and minimum

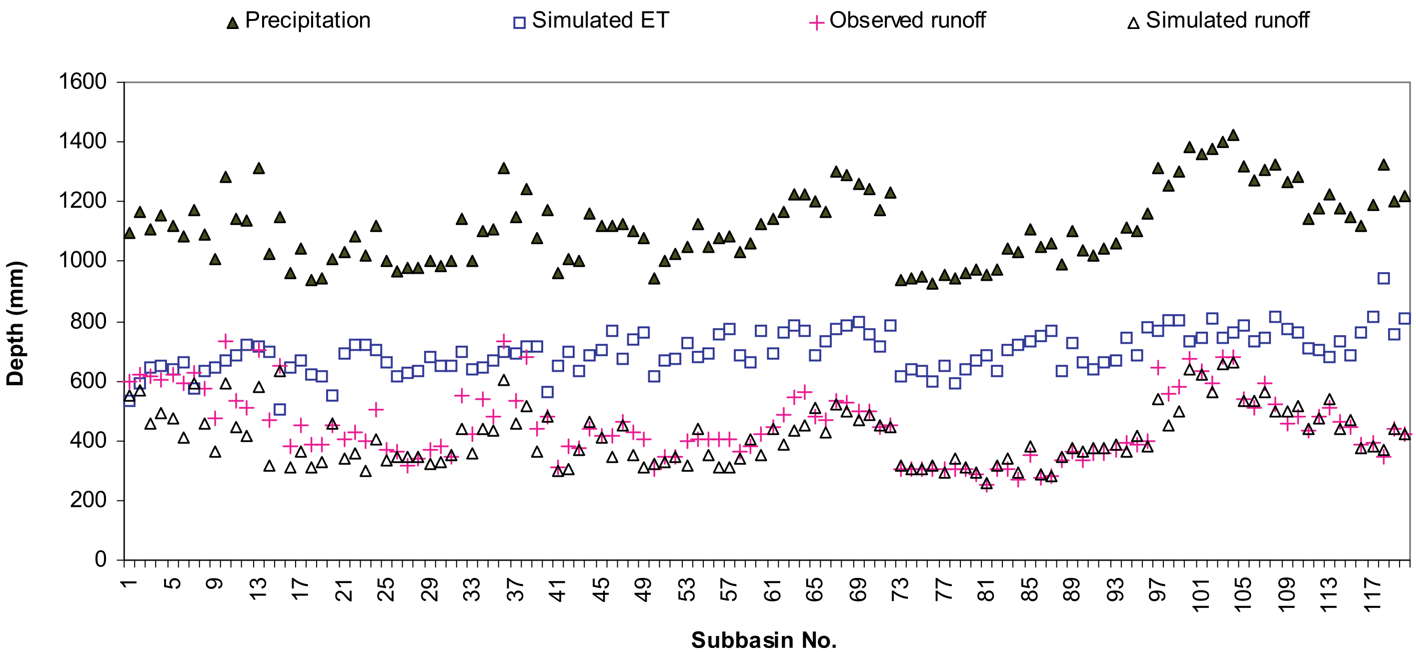

FIGURE 2. Precipitation, Simulated ET, Simulated Runoff, and Observed Runoff for the Eight-Digit Watersheds in the Ohio River Basin.

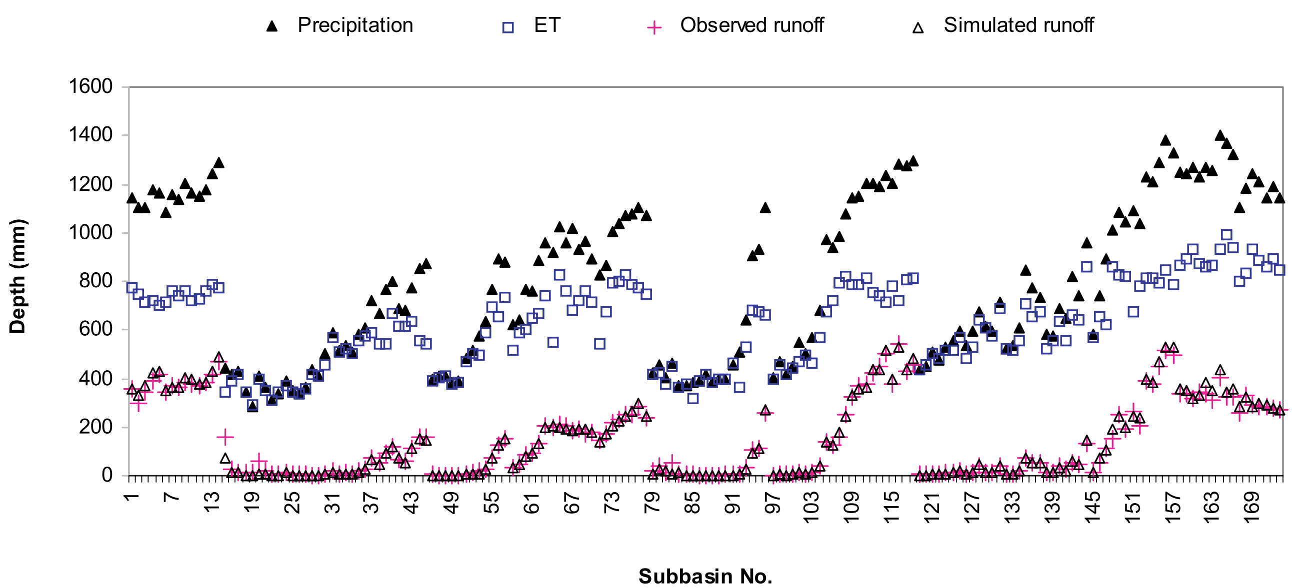

FIGURE 3. Precipitation, Simulated ET, Simulated Runoff and Observed Runoff for the Eight-Digit Watersheds in the Arkansas-White-Red River Basin.

temperature and radiation, which can vary spatially. Figure 4 shows the effects of HARGO_PETCO on simulated long-term annual average runoff in various subwatersheds (eight-digit) that were calibrated in the Ohio River Basin. Not all eight-digit watersheds required calibration. Runoff increased in most of the subwatersheds during calibration by reducing HARG_PETCO to reduce PET. The magnitude of increase in runoff varied across subwatersheds (shown by bars in top, that is, difference in runoff before and after adjustment of HARG_PETCO). Hargreaves method accounts for variations in minimum and maximum temperature, and variations in crops are accounted simultaneously while estimating the actual ET.

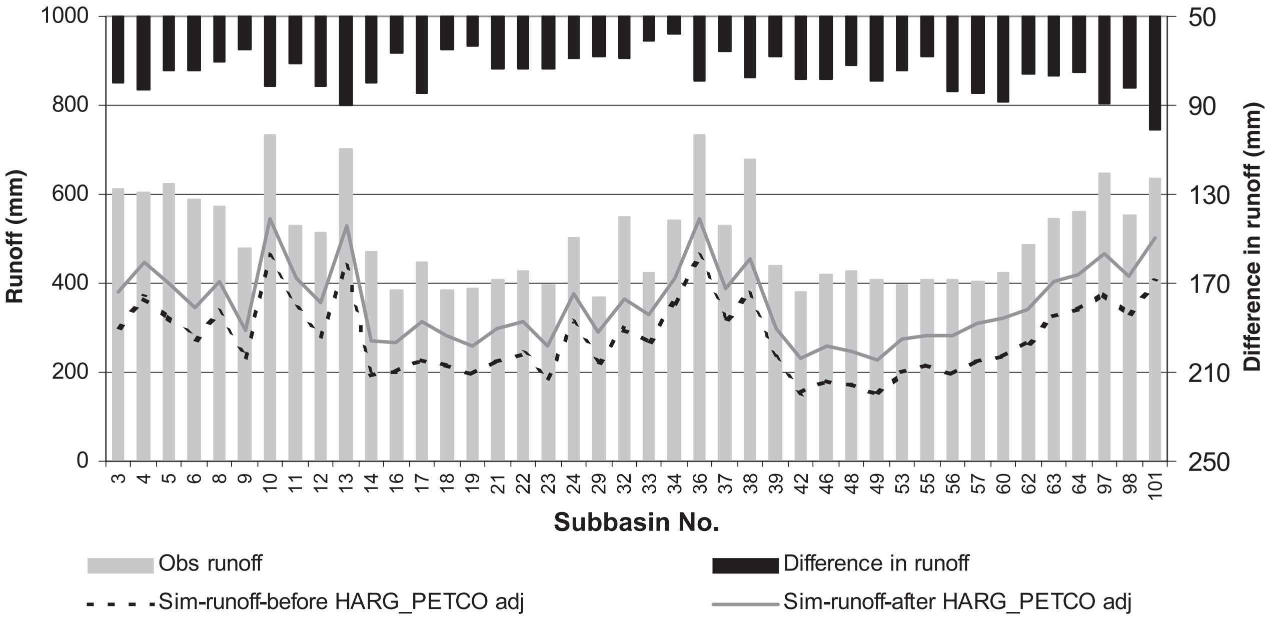

FIGURE 4. Effect of HARG_PETCO on Spatial Variation in Simulated Runoff in the Eight-Digit Watersheds That Were Calibrated in the Ohio River Basin.

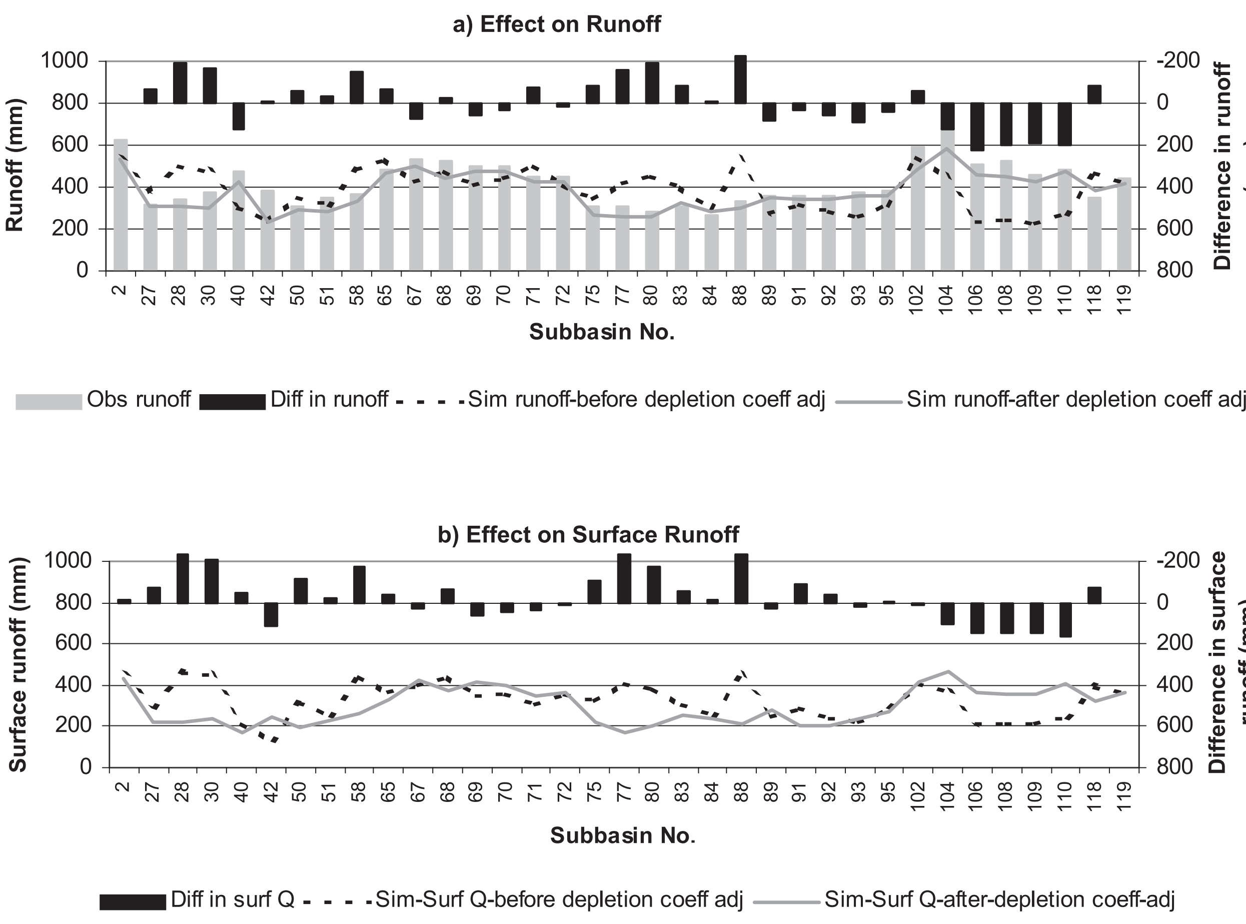

FIGURE 5. Effect of Soil Water Depletion Coefficient on Spatial Variation in Simulated Runoff and Surface Runoff in the Eight-Digit Watersheds That Were Calibrated in the Ohio River Basin (note: upward bars—decrease and downward bars—increase in runoff and surface runoff after soil water depletion coefficient adjustment).

2. Soil water depletion coefficient (CN_COEF) is used for calibrating surface runoff in SWAT (Nietsch *et al.*, 2002; Kannan *et al.*, 2007b). Wang *et al.* (2006) indicated that surface runoff is sensitive to the soil water depletion coefficient. It can be used to calibrate surface runoff and ground water proportions. The model overestimated and underestimated surface runoff in certain eight-digit watersheds. Figures 5a and 5b show the effects of the soil water depletion coefficient on simulated runoff and surface runoff in various subwatersheds in the Ohio River Basin. Magnitudes of the increase or decrease in runoff and surface runoff varied across subwatersheds (shown by upward and downward bars) with changes in soil water depletion coefficient. It could be noticed that changes made in surface runoff and runoff showed the trend to match the observed or targeted runoff.
3. The spatial effects of ground-water re-evaporation coefficient (GWREVAP) on long-term average annual runoff and groundwater are shown in Figures 6a and 6b. It could be noticed that the changes observed in runoff were due to changes in ground water in response to the GWREVAP coefficient. The changes were up to 15 mm.

The other model input parameters were adjusted in the similar manner for runoff calibration in each eight-digit watershed.

The average annual simulated runoff and average annual observed runoff of the eight-digit watersheds in the Ohio region and Arkansas region are shown in Figures 7 and 8, respectively. In the Ohio region, the observed runoff increased from northwest to the southeast similar to precipitation pattern (Figure 7). The simulated runoff showed similar pattern by capturing the spatial variations in runoff across the region (Figure 7). The regression relationship between observed and simulated runoff at eight-digit watersheds indicate that the model prediction is satisfactory (Figure 9a). Out of 120 hydrologic unit codes (HUCs) in the basin, the simulated runoff in 106

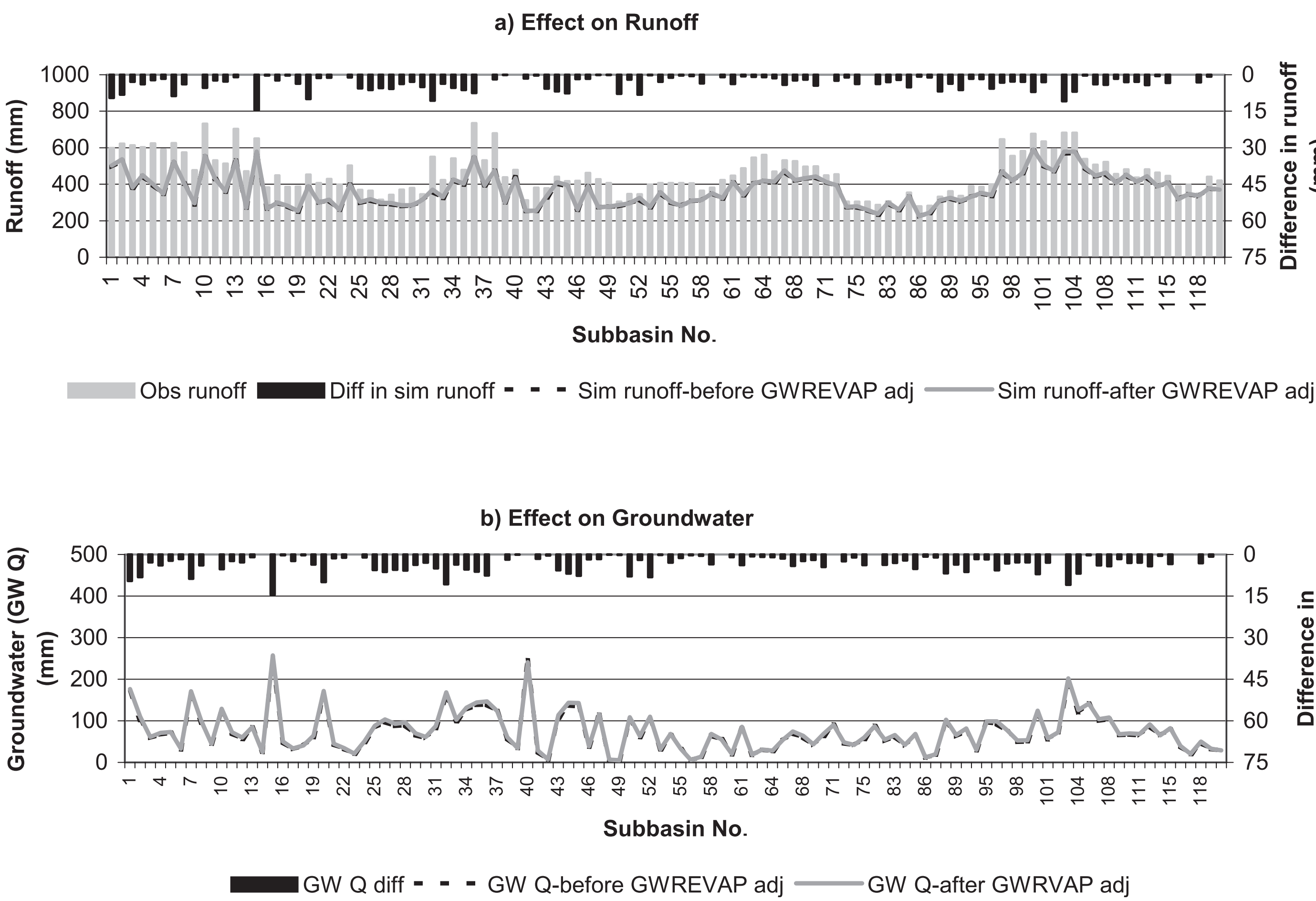

FIGURE 6. Effect of Ground-Water Revap Coefficient on Spatial Variation in Simulated Runoff and Ground Water in the Eight-Digit Watersheds That Were Calibrated in the Ohio River Basin.

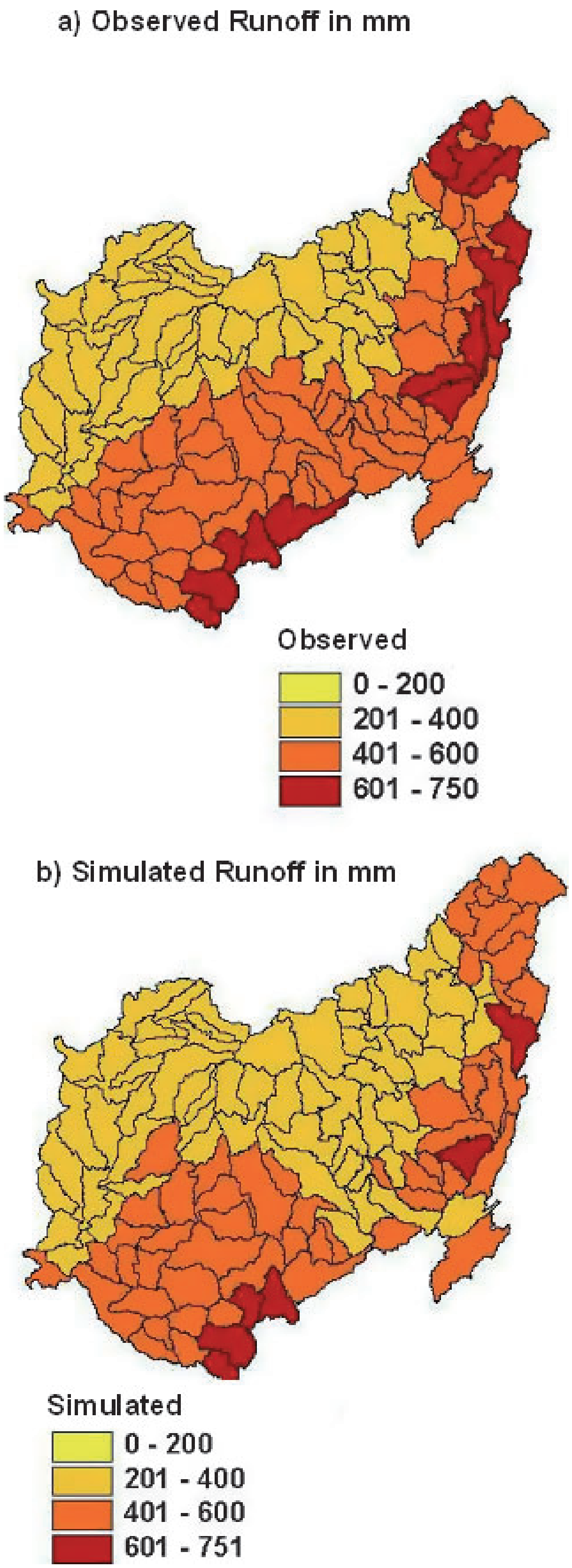

FIGURE 7. Observed and Simulated Average Annual Runoff for Eight-Digit Watersheds in the Ohio River Basin.

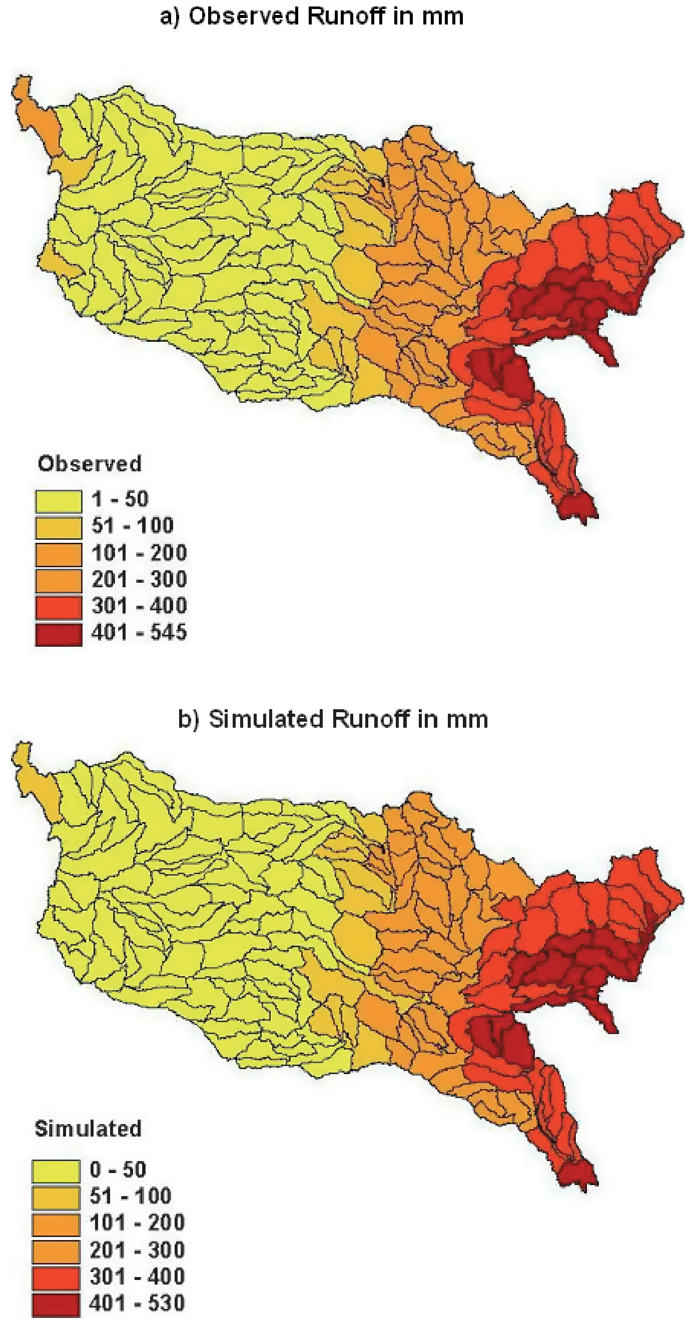

FIGURE 8. Observed and Simulated Average Annual Runoff for Eight-Digit Watersheds in the Arkansas-White-Red River Basin.

HUCs were within 20% of the observed runoff. There were underpredictions of runoff in a few HUCs where the runoff was high in the range of 600-700 mm (Figure 7). Further investigation showed that the model underpredicted the base flow portion in those HUCs that were not matching the calibration criteria. Snowfalls and snowmelting are a common phenomena in the Ohio region and the model had difficulties dealing with it.

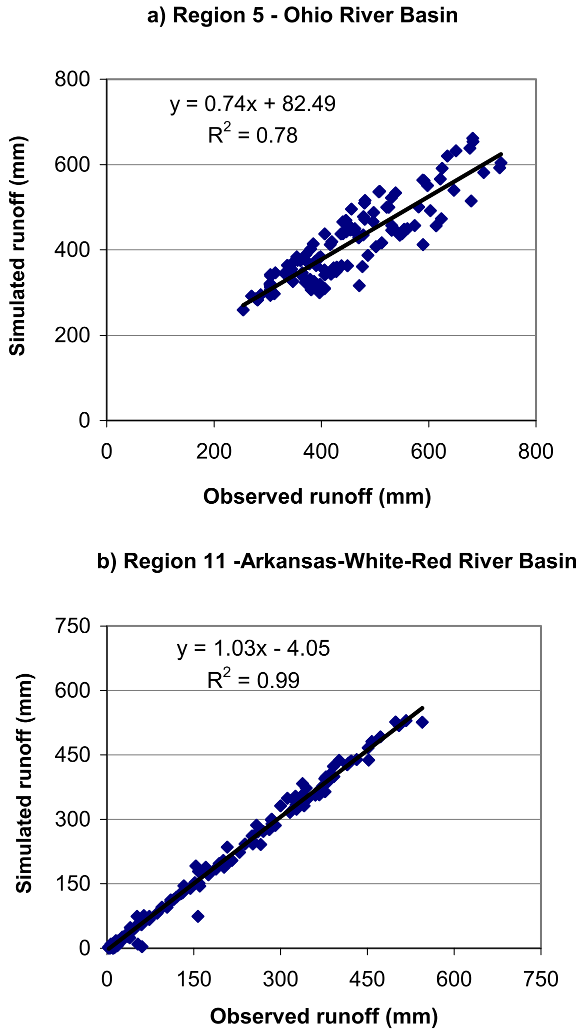

FIGURE 9. Regression Relationship Between Average Annual Observed and Simulated Runoff at Eight-Digit Watersheds in the Ohio and Arkansas-White-Red River Basins.

In the Arkansas-White-Red River region, the average annual observed runoff varied widely from <50 mm in the western side through more than 500 mm in the eastern side. Simulated runoff matched this spatial variation pattern very well (Figure 8). Observed and simulated runoff patterns are in concurrence with the precipitation patterns of this region (Figure 3). The regression coefficient of 0.99 (Figure 9b) revealed that the observed and simulated runoff matched very well at eight-digit watersheds in this region. The simulated runoff was within 20% of the observed runoff in 128 HUCs out of 173 HUCs in this region. As runoff is relatively low in majority of the HUCs in this region (Figure 8), considering the absolute difference in runoff would be a better indication than percentage difference. The simulated runoff was within 25 mm in 159 HUCs and within of 50 mm in 171 HUCs, when compared to the observed runoff (Figure 8).

Results of the two study regions indicate that the SWAT model is able to capture the spatial variations in runoff and simulate the local water balances adequately.

Temporal Validation of Streamflow at Multiple Gaging Locations in the Main River

Without further calibration, regression of observed and simulated annual and monthly streamflow was performed to validate the model.

Ohio River Basin: The observed and simulated annual and monthly streamflows on the Ohio River at Louisville, Kentucky (USGS Station 03294500) and Metropolis, IL (USGS Station 03294500) matched well (Figure 10 and 11). Means of the observed and simulated annual and monthly flows were within a difference of 10% at Louisville, Kentucky (Table 3). Further agreement between annual and monthly simulated and observed flows at Louisville are shown by the coefficient of determination >0.6 and NSE >0.5 (Table 3). Good agreement between annual and monthly observed, and simulated flows at Metropolis, Illinois, is indicated by the time series plots and statistics (Figure 11 and Table 3). However, there is a general tendency for the model to underpredict the peak flows during spring months and sometimes overpredict the base flow during fall months. This may be due to either limitations in snowmelt simulation or simulation of the reservoir operations.

The Arkansas-White-Red River Basin: This is relatively a low flow region. The observed and simulated annual and monthly streamflows along the Arkansas River at Arkansas City, Kansas (USGS Station 07146500) matched moderately well except for overprediction of peak flows (Figure 12) in a few years including 1973. As NSE is sensitive to outliers, the NSE computed was low because of the overestimation of peak flows. Further investigations revealed that there were major rainfall events during the months of March and October in 1973 and the model overpredicted the runoff events. Similarly, there was a consistent underprediction of the peaks during May/June in most of the years. The model was not able to simulate the sudden changes in flow variations as seen in the observed flow. Hence, the simulated annual average flows were lower.

It could be observed from Figure 13a and statistics shown in Table 3 that the observed and simulated annual flows compared fairly well at Index on the Red River, Arkansas (USGS Station 07337000). Simulated monthly flows were closer to the observed flows at this location (Figure 13b and Table 3).

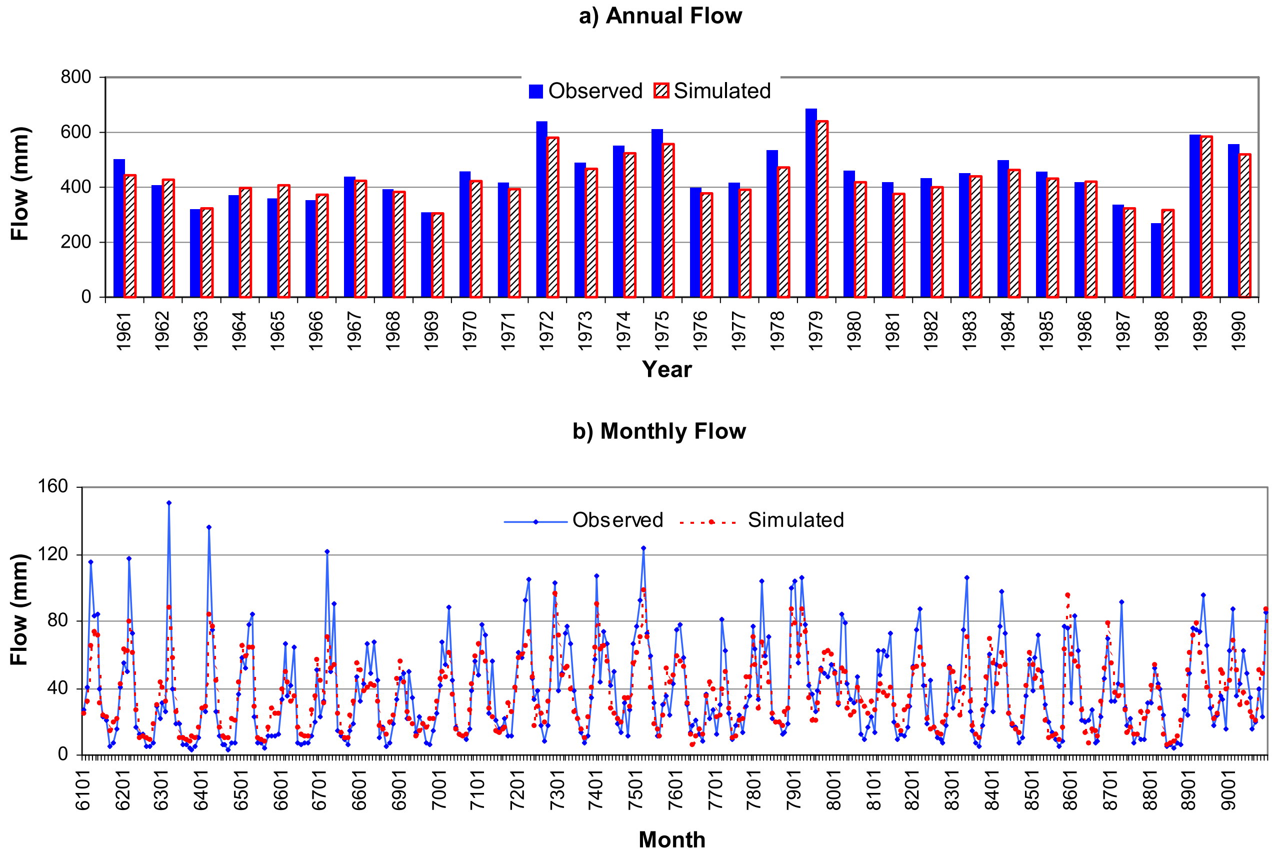

FIGURE 10. Annual and Monthly Observed and Simulated Flows at Louisville, Kentucky, on the Ohio River.

Overall, the model preserved the peaks and recessions. In this region also, there is a general tendency for the model to underpredict peaks during spring months.

It should be noticed that the mean annual flow varied widely between the two regions (Figures 10-13). The mean annual streamflow at the gages analyzed in the Ohio River Basin were approximately 450 mm and while it varied from 15 to 90 mm in the Arkansas-White-Red River Region. Mean annual and monthly flows at the two gaging stations in the Ohio River Basin were in the similar ranges. However, in the case of Arkansas-White-Red River Basin, there were variations in mean annual flow between the gages at Arkansas City on the Arkansas River and at Index on the Red river. The time series annual and monthly flow results at the gages in both the regions appeared to be reasonable given that no additional calibration was performed after the spatial runoff calibration at eight-digit watersheds. Overall, the model is able to capture the annual and monthly flow patterns.

Results of runoff calibration at eight-digit watersheds and streamflows at gaging stations indicate that the hydrological variations at spatial and temporal scales are simulated reasonably well. This study has shown the importance of a spatial calibration along with temporal validation, especially when there is a wide variation in runoff across the basin. Watershed characteristics within and between subwatersheds differ in terms of precipitation, other weather parameters, land use and land cover, topography, soils, and crops grown. These watershed characteristics generate variable hydrologic patterns across the river basin. The calibration and validation approach needs to capture the variations in flow patterns at subwatershed and watershed level for reliable simulations of water flow. Reasonable accuracy in flow simulation is necessary for simulating the transport of pollutants. Once the flow is estimated reasonably well, the model can be calibrated and validated for sediment and nutrients and can be used for several applications, including (1) identification of subwatersheds that have critical sediment/erosion problems and, (2) identification of subwatersheds or watershed region that contribute excessive nitrogen and phosphorus loadings to the river system, and (3) estimation of benefits of conservation practices on water quality in terms of percentage reductions in sediment, nutrients, and pesticide loadings. The

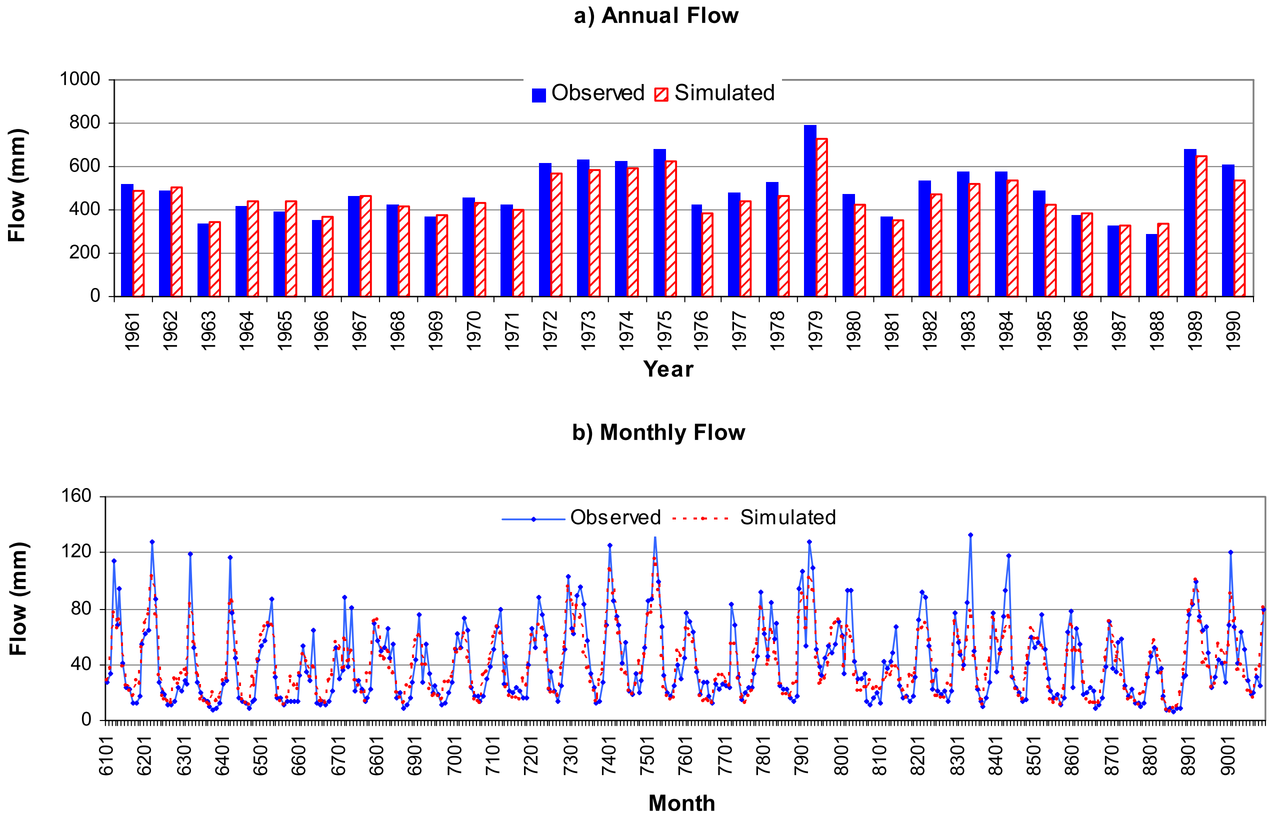

FIGURE 11. Annual and Monthly Observed and Simulated Flows at Metropolis, Illinois, on the Ohio River.

model can be also used to predict concentrations of nitrogen and phosphorus in the river systems to meet water quality standards for humans and eco-systems and identify the sources of excessive nutrient contributions.

SUMMARY AND CONCLUSIONS

Physically based regional scale hydrologic modeling is useful in investigating the effects of different management scenarios on water quality and quantity. Calibration and validation of the model for the study region are necessary to capture the variable hydrological patterns in subwatersheds and watershed. This is especially important in large river basins with wide spatial and temporal variations in flow patterns. In addition, availability of limited observed data for model validation makes the regional scale study challenging. In this study, regional scale hydrologic modeling is described for two river basins, and a flow calibration/validation procedure involving calibration of spatial variation of annual average runoff at subwatershed level (to assure local water balance), and validation of the time series of flow at key locations along the main river (to assure temporal variability) is carried out using the SWAT. The regional scale modeling procedure is demonstrated with results from two river basins, the Ohio and Arkansas-White-Red River basins that are in different hydrologic conditions. The long-term average annual runoff estimated from the USGS data for the eight-digit watersheds were used for conducting the spatially distributed calibration. R^2 values of average annual runoff at subwatersheds were 0.78 and 0.99 for the Ohio and Arkansas Basins. The annual and monthly streamflow data from the USGS gages from 1961-1990 were used for temporal flow validation. R^2 values of the annual and monthly flows for the multiple gaging stations studied at Ohio and Arkansas were >0.6. It is expected that the calibration and validation approach similar to this study would improve the reliability of hydrologic model predictions at regional scale river basins. Because of the large-scale nature of the study and limitation in availability of time series of observed data, average annual runoff was used for spatial calibration. The average annual runoff value is a good indicator of water balance in a subwatershed and this approach seemed to provide realistic prediction of the annual

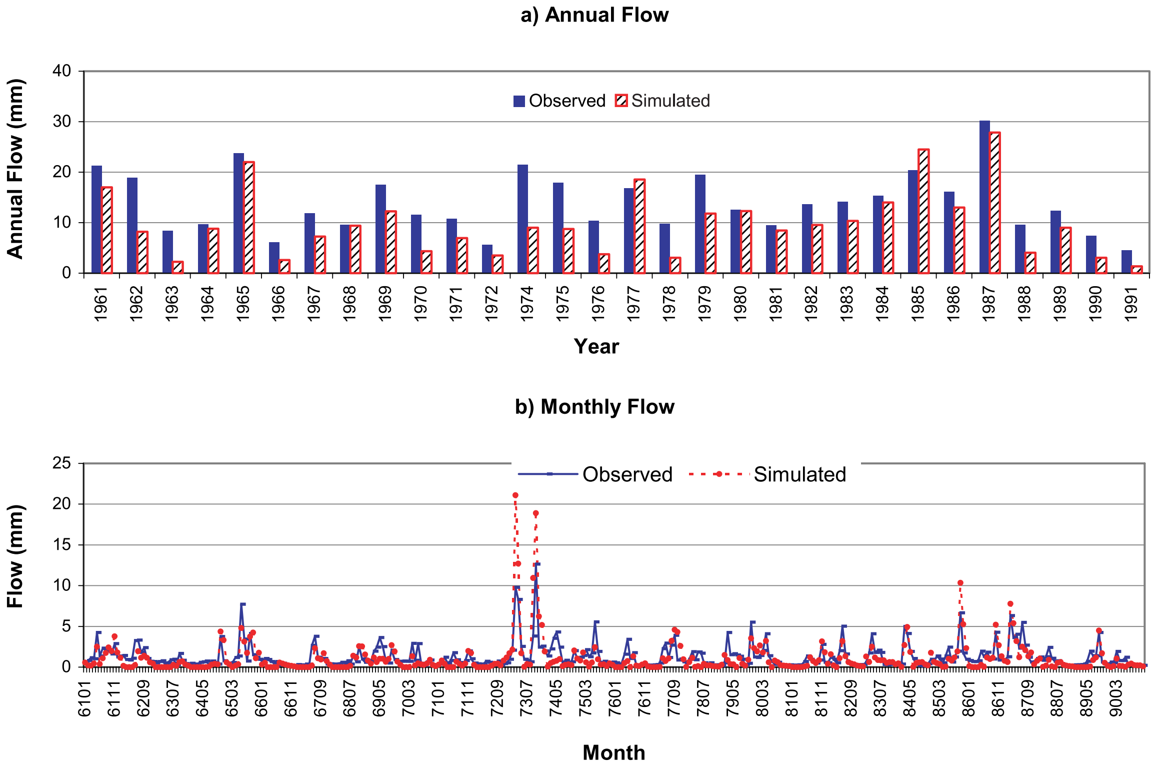

FIGURE 12. Annual and Monthly Observed and Simulated Flows on the Red River at Index, Arkansas, in the Arkansas-White-Red River Basin.

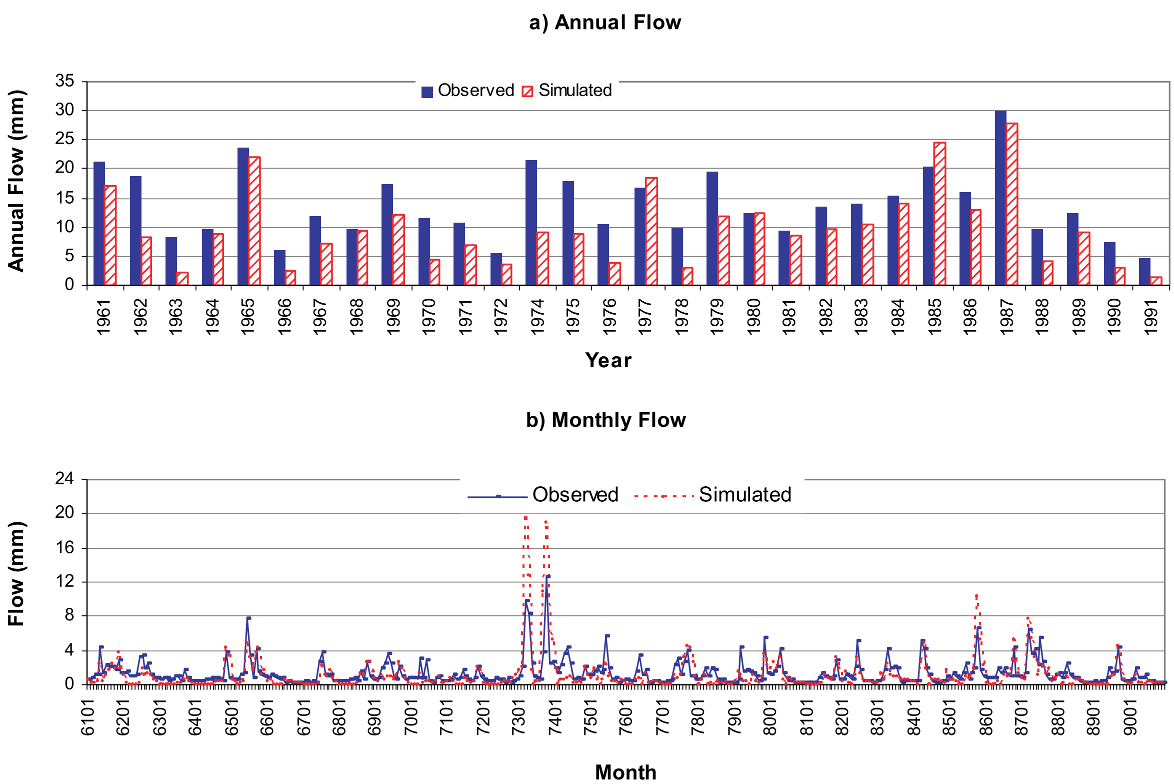

FIGURE 13. Annual and Monthly Observed and Simulated Flows on the Arkansas River at Arkansas City, Kansas, in the Arkansas-White-Red River Basin.

and monthly flow pattern at multiple locations along the main river.

Major conclusions from this study include

(1) Compared to the traditional approach of calibrating and validating at the watershed outlet, it is expected that the spatial calibration and validation approach would improve the reliability of hydrologic model predictions by capturing the variations in flow patterns at subwatershed and watershed levels for large/regional scale river basins.
(2) When tested in two river basins, the spatial calibration process seems to be helpful in capturing the flow variations from low flow through high flow regimes. Reasonably accurate prediction of flow is a pre-requisite for reliable predictions of sediment and nutrient yields.
(3) The application of spatial calibration and temporal validation approach to large-scale studies can be demonstrated with CEAP and/or other agricultural management and water quality projects.
(4) Current regional scale modeling framework can be used for potential applications such as to assess the effects of land use changes on water quality and quantity, and assess the effects of climate changes on water budget at regional scale. The modeling framework can also be used by planners and managers to address several policy-related questions on water supply and water quality management issues.

ACKNOWLEDGMENTS

The USDA-NRCS Resource Inventory Assessment Division provided funding for this work as part of the Conservation Effects Assessment Project (CEAP). Thanks to the editor and the anonymous reviewers for their constructive comments. The Agricultural Policy/Environmental extender (APEX) modeling team's contribution is acknowledged.

LITERATURE CITED

Arnold, J.G., R.S. Muttiah, R. Srinivasan, and P.M. Allen, 2000. Regional Estimation of Baseflow and Groundwater Recharge in the Upper Mississippi River Basin. Journal of Hydrology 227:21-40.

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and P.M. Allen, 1999. Continental Scale Simulation of the Hydrologic Balance. Journal of the American Water Resources Association 35(5):1037-1051.

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams, 1998. Large Area Hydrologic Modeling and Assessment Part I: Model Development. Journal of the American Water Resources Association 34(1):73-89.

Borah, D.K. and M. Bera, 2004. Watershed Scale Hydrologic and Nonpoint Source Pollution Models: Review of Applications. Transactions of the American Society of Agricultural Engineers 47(3):789-803.

Daly, C., W.P. Gibson, G.H. Taylor, G.L. Johnson, and P. Pasteris, 2002. A Knowledge Based Approach to the Statistical Mapping of Climate. Climate Research 22:99-113.

Di Luzio, M., G.L. Johnson, C. Daly, J. Eischeid, and J.G. Arnold, 2008. Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. Journal of Applied Meteorology and Climatology 47:475-497.

Eischeid, J.K., P.A. Pasteris, H.F. Diaz, M.S. Plantico, and N.J. Lott, 2000. Creating a Serially Complete, National Daily Time Series of Temperature and Precipitation for the Western United States. Journal of Applied Meteorology 39:1580-1591.

Gassman, P.W., M.R. Reyes, C.H. Green, and J.G. Arnold, 2007. The Soil and Water Assessment Tool: Historical Development, Applications and Future Research Directions. Transactions of the American Society of Agricultural and Biological Engineers 50(4):1211-1250.

Gebert, W.A., D.J. Graczyk, and W.R. Krug, 1987. Average Annual Runoff in the United States, 1951-1980. Hydrologic Investigations Atlas, HA-70, U.S. Geological Survey, Reston, Virginia.

Hao, F.H., X.S. Zhang, and Z.F. Yang, 2004. A Distributed Nonpoint Source Pollution Model: Calibration and Validation in the Yellow River Basin. Journal of Environmental Sciences 16(4):646-650.

Hargreaves, G.L. and R.G. Allen, 2003. History and Evaluation of Hargreaves Evapotranspiration Equation. Journal of Irrigation and Drainage Engineering 129(1):53-63.

Hargreaves, G.L., G.H. Hargreaves, and J.P. Riley, 1985. Agricultural Benefits for Senegal River Basin. Journal of Irrigation and Drainage Engineering 111(2):113-124.

Hargreaves, G.H. and Z.A. Samani, 1985. Reference Crop Evapotranspiration From Temperature. Applied Engineering in Agriculture 1:96-99.

Hitt, K.J., 1985. Surface Water and Related Land Resources Development in the United States and Puerto Rico: U.S. Geological Survey Special Map, Scale 1:3,168,000.

Jha, M., J.G. Arnold, P.W. Gassman, F. Giorgi, and R. Gu, 2006. Climate Change Sensitivity Assessment on Upper Mississippi River Basin Streamflows Using SWAT. Journal of the American Water Resources Association 42(4):997-1015.

Kannan, N., C. Santhi, and J.G. Arnold, 2008. Development of an Automated Procedure for estimation of the spatial variation of runoff in large river basins. Journal of Hydrology (In review after revision).

Kannan, N., C. Santhi, J.R. Williams, and J.G. Arnold, 2007. Development of a Continuous Soil Moisture Accounting Procedure for Curve Number Methodology and its Behaviour With Different Evapotranspiration Methods. Hydrological Processes, doi: 10.1002/hyp 6811.

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, and T.L. Veith, 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the American Society of Agricultural and Biological Engineers 50(3):885-900.

Nash, J.E. and J.V. Suttcliffe, 1970. River Flow Forecasting Through Conceptual Models, Part I—A Discussion of Principles. Journal of Hydrology 10(3):282-290.

Neitsch, S.L., J.G. Arnold, J.R. Williams, J.R. Kiniry, and K.W. King, 2002. Soil and Water Assessment Tool (Version 2000)—Theoretical Documentation. GSWRL 02-01, BREC 02-05, TR-191.: Texas Water Research Institute, College Station, Texas.

Nicks, A.D., 1974. Stochastic Generation of the Occurrence, Pattern and Location of Maximum Amount of Rainfall. *In* Proceedings of Symposium on Statistical Hydrology, Tuscon, Arizona, Aug-Sep 1971, USDA, Misc. Publ. No. 1275. pp. 154-171.

Qi, C. and S. Grunwald, 2005. GIS-Based Hydrologic Modeling in the Sandusky Watershed Using SWAT. Transactions of the American Society of Agricultural and Biological Engineering 48(1):169-180.

Santhi, C., J.G. Arnold, J.R. Williams, W.A. Dugas, R. Srinivasan, and L.M. Hauck, 2001. Validation of the SWAT Model on a Large River Basin With Point and Nonpoint Sources. Journal of the American Water Resources Association 37(5):1169-1188.

Santhi, C., N. Kannan, M. Di Luzio, S.R. Potter, J.G. Arnold, J.D. Atwood, and R.L. Kellogg, 2005. An Approach for Estimating Water Quality Benefits of Conservation Practices at the National Level. ASAE 2005 Annual Meeting, Tampa, Florida. Paper No. 052043.

Sharpley, A.N. and J.R. Williams (Editors). 1990. EPIC—Erosion Productivity Impact Calculator, Model Documentation, Tech. Washington, D.C. Bulletin No. 1768, USDA-ARS. p. 235.

Smith, R.A., G.E. Schwarz, and R.B. Alexander, 1997. Regional Interpretation of Water Quality Monitoring Data. Water Resources Research 33:2781-2798.

Srinivasan, R., J.G. Arnold, and C.A. Jones, 1998. Hydrologic Unit Modeling of the United States With the Soil and Water Assessment Tool. International Journal of Water Resources Development 14(3):315-325.

U.S. Army Corps of Engineers, 1982. National Inventory of Dams Database in Card Format (Computer Tape). Available from National Technical Information Service, Springfield, Virginia. #ADA 118670.

USDA-NRCS, 1994. State Soil Geographic Database. United States Department of Agriculture-Natural Resources Conservation Service. http://www.soils.usda.gov/survey/geography/statsgo/, *accessed* November 2006.

USDA-SCS, 1972. National Engineering Handbook. USDA-Soil Conservation Service, Washington, D.C. Chaps. 4-10.

USEPA (U.S. Environmental Protection Agency), 1998. Water Pollution Control: 25 Years of Progress and Challenges for the New Millennium. EPA 833-F-98-003. USEPA Office of Wastewater Management, Washington, D.C.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel, 2001. Completion of the 1990s National Land Cover Data Set for the Conterminous United States From Landsat Thematic Mapper Data and Ancillary Data Sources. Photogrammetric Engineering and Remote Sensing 67:650-652.

Wang, X., S.R. Potter, J.R. Williams, J.D. Atwood, and T. Pitts, 2006. Sensitivity Analysis of APEX for National Assessment. Transactions of the American Society of Agricultural and Biological Engineering 49(3):679-688.

Wolock, D.M. and G.J. Mc Cabe, 1999. Explaining Spatial Variability in Mean Annual Runoff in the Conterminous United States. Climate Research 11:149-159.